Blackford: A Dual Processor Chipset for Servers and Workstations

Kai Cheng, Sundaram Chinthamani, Sivakumar Radhakrishnan, Fayé Briggs and Kathy Debnath

Intel Corporation

8/22/2006



© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Legal Disclaimer

- Intel, the Intel logo, Centrino, the Centrino logo, Intel Core, Core Inside, Pentium, Pentium Inside, Itanium, Itanium Inside, Xeon, Xeon Inside, Pentium III Xeon, Celeron, Celeron Inside, and Intel SpeedStep are trademarks or registered trademark of Intel Corporation or its subsidiaries in the United States and other countries.
- This document is provided “as is” with no warranties whatsoever, including any warranty of merchantability, non-infringement fitness for any particular purpose, or any warranty otherwise arising out of any proposal, specification or sample
- Information in this document is provided in connection with Intel products. No license, express or implied, by estoppels or otherwise, to any intellectual property rights is granted by this document. Except as provided in Intel's Terms and Conditions of Sale for such products, Intel assumes no liability whatsoever, and Intel disclaims any express or implied warranty, relating to sale and/or use of Intel products including liability or warranties relating to fitness for a particular purpose, merchantability, or infringement of any patent, copyright or other intellectual property right. Intel products are not intended for use in medical, life saving, or life sustaining applications.
- Intel does not control or audit the design or implementation of 3rd party benchmarks or websites referenced in this document. Intel encourages all of its customers to visit the referenced websites or others where similar performance benchmarks are reported and confirm whether the referenced benchmarks are accurate and reflect performance of systems available for purchase.
- Intel may make changes to specifications and product descriptions at any time, without notice.
- All plans, features and dates are preliminary and subject to change without notice.
- **Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Buyers should consult other sources of information to evaluate the performance of systems or components they are considering** purchasing.
- * Third-party brands and names are the property of their respective owners.
- Copyright © Intel Corporation 2006

© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Agenda

- Bensley Platform Overview
- Blackford North Bridge Micro-Architectural Features
- Performance
- Summary



© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Volume DP Server/Workstation Platform

New Platform Bus Architecture

- Dual Independent Front Side buses
 - Dempsey (Xeon 5000 series) / Woodcrest (Xeon 5100 series) dual core and the upcoming quad core Xeon processor support
 - Faster FSB speeds, higher throughput, better performance (10.6GB/s theoretical limit)
 - Lower board costs with improved routing

Memory

- 4 FBD channels @ 21 GB/s (theoretical limit) for reads and 10.6 GB/s (theoretical limit) for writes
- DDR2-533/667 FBD, 16 dimm maximum

Comprehensive RAS Features

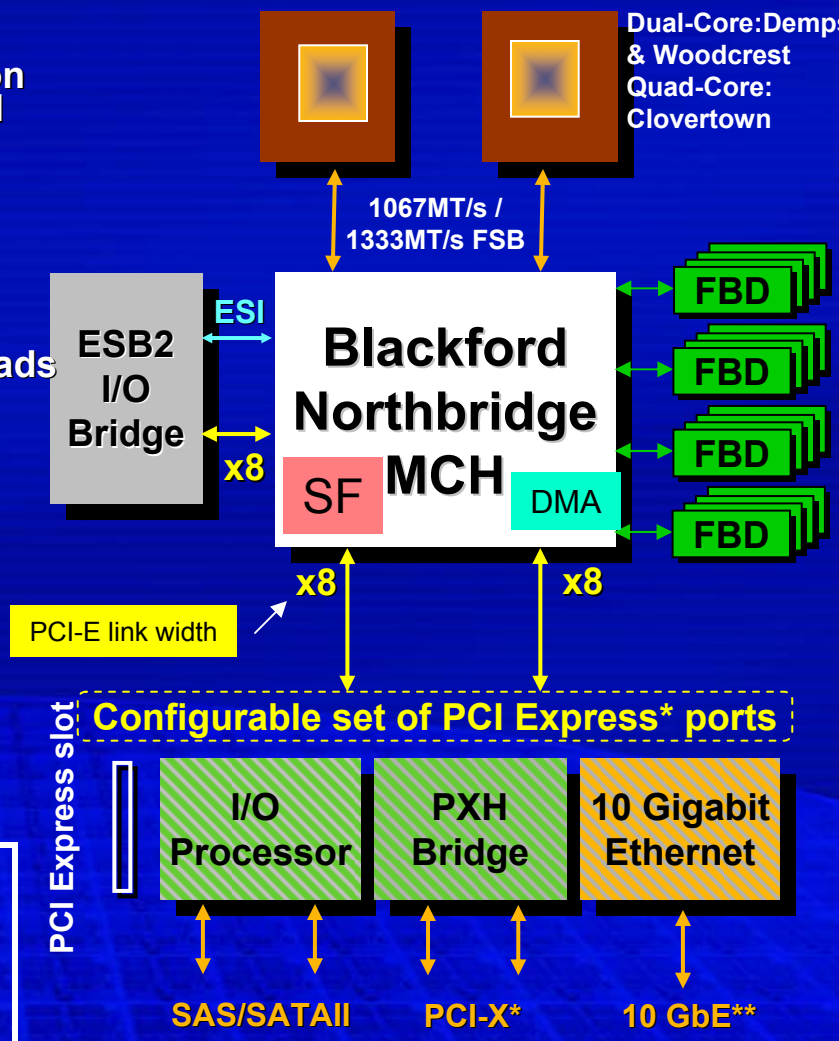
Direct connect I/O

- 3 x8 PCI Express and x4 ESI links @ 2.5GT/s
- Enterprise South Bridge other peripheral devices
- Integrated Direct Memory Access (DMA) support

Greencreek Workstation Configuration

- Fully inclusive Snoop Filter (SF) to track 16MB effective cache coverage

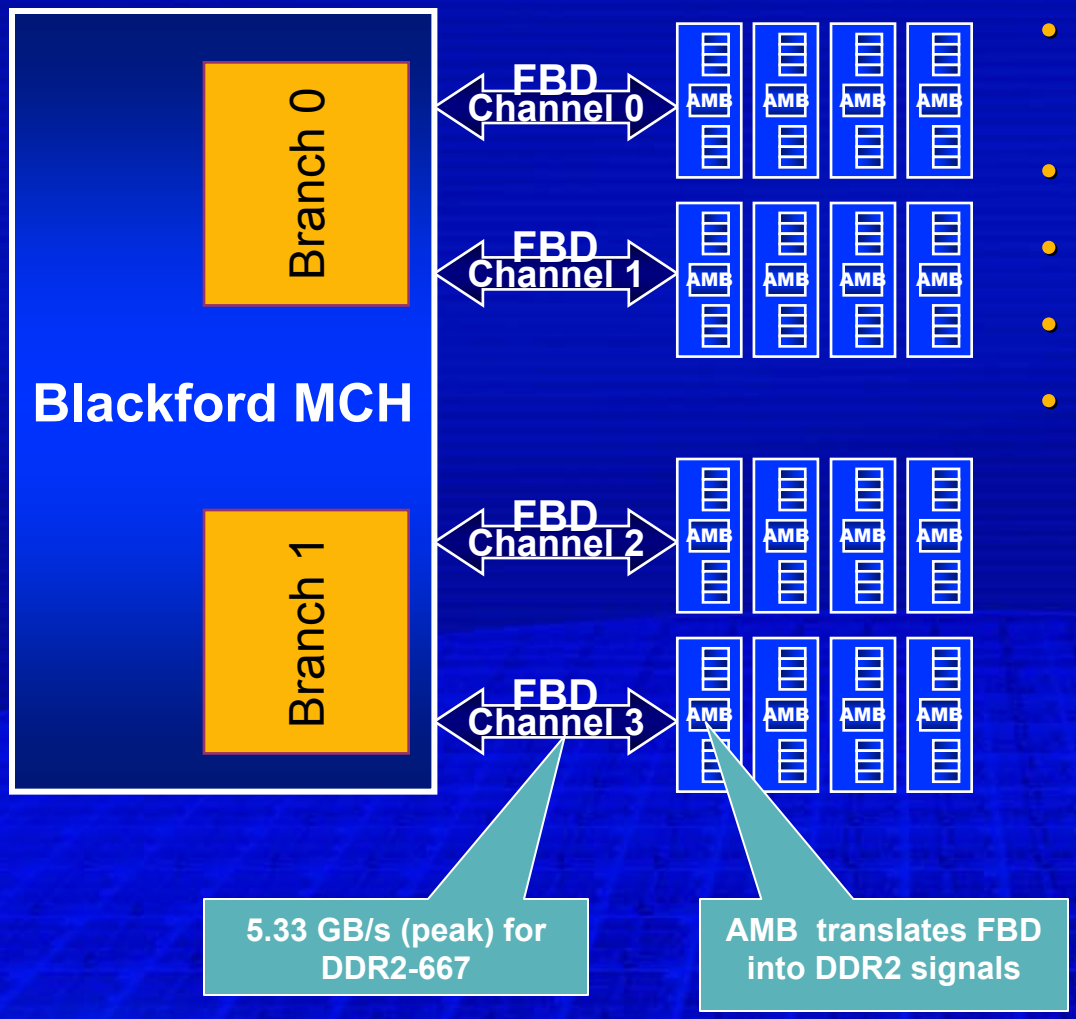
Process: 130nm, 6 Metal layer CMOS
Package Size: 42.5mm 1432 balls at 1.092mm ball pitch
Transistor count: ~52M (Blackford); ~65M (Greencreek)
TDP Power: ~27w (Blackford); ~32.4w (Greencreek)



Feature & performance leadership in DP

© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

FBD Memory Interface



- 4 Fully Buffered DIMM channels
- 2 Channels per branch
- 8 ranks per branch.
- 16 DIMMs max
- Branch Lockstep mode of operation
 - 64B Cache-line split across a branch i.e. 32B per channel
 - Burst length of 4 used on DDR2 memory

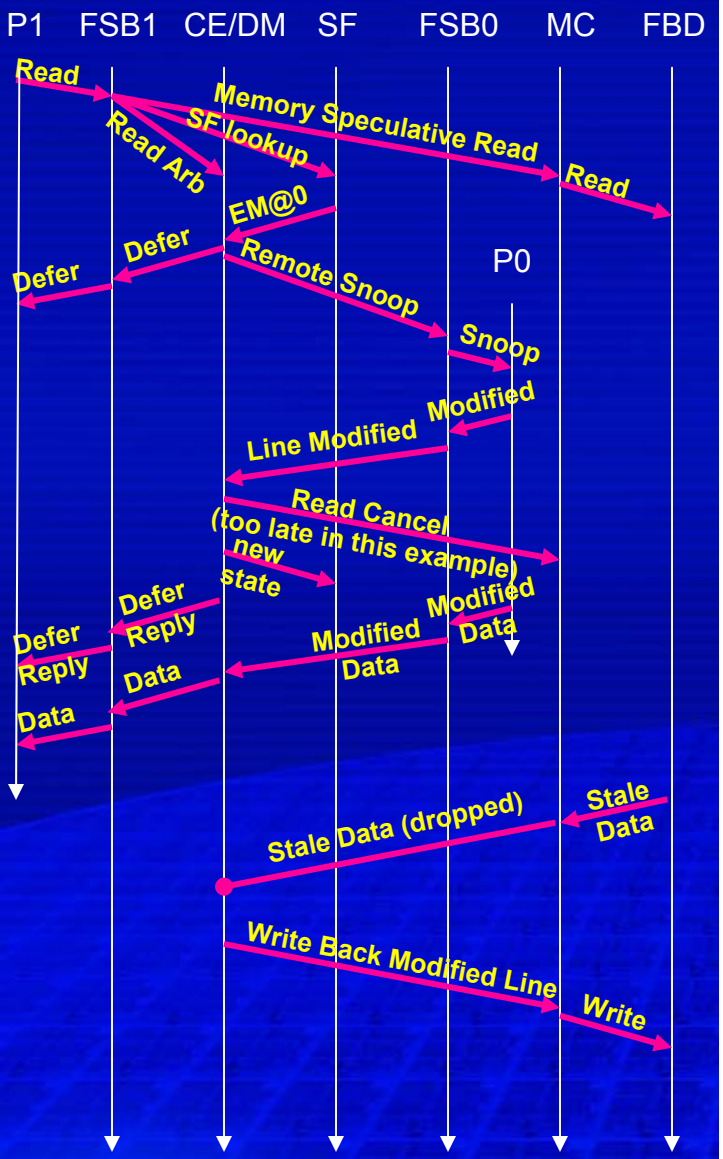
© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

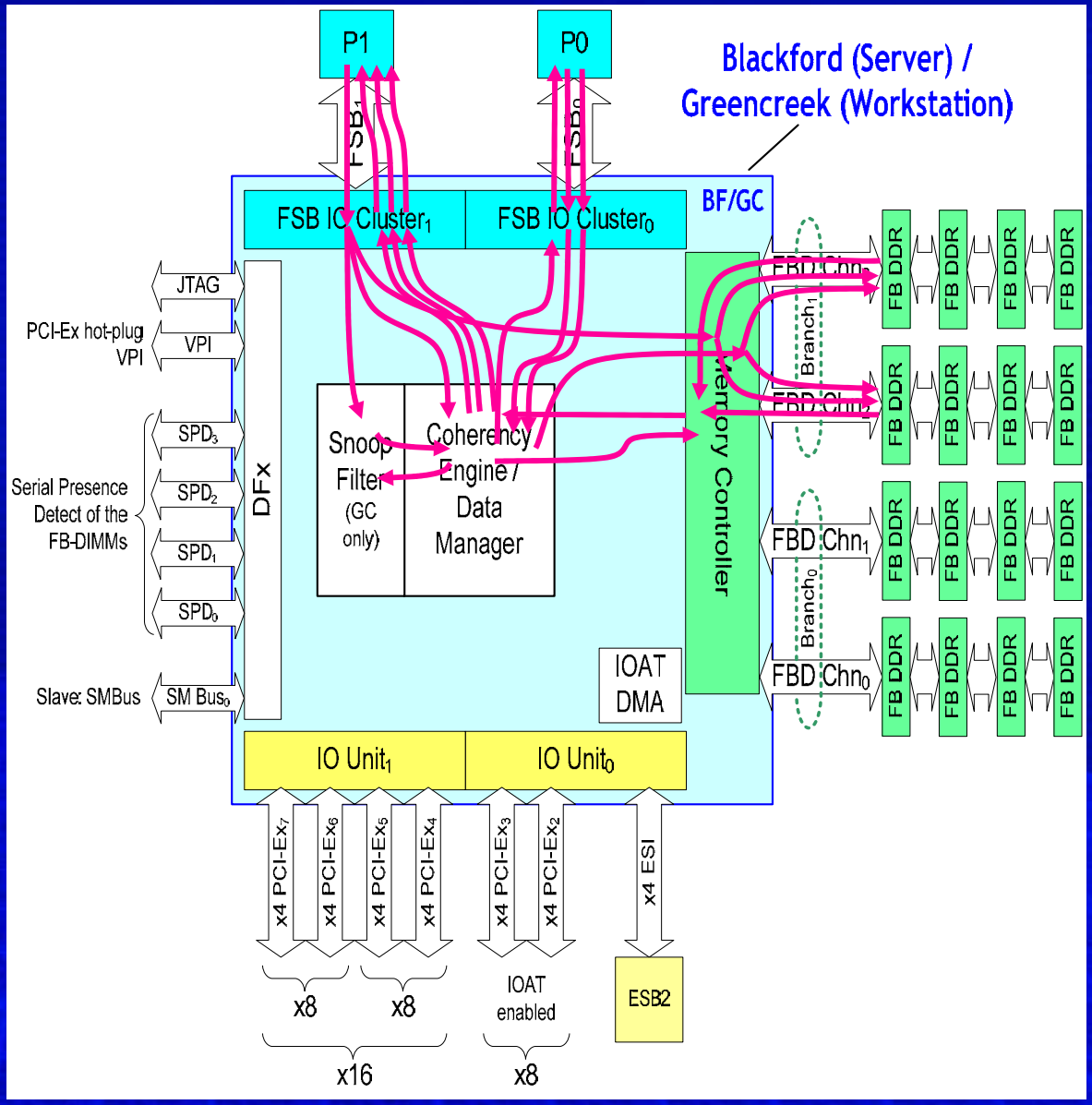
Enterprise Server Technologies

Technology	Benefits
Low power Dual Core with Intel® 64 Technology	Seamless 32/64-bit computing, >4GB of Physical address space
Intel® I/O Acceleration Technology	Offloads CPU for memory copy through on-die DMA engine
Intel® Virtualization Technology	Enables multiple Operating Systems to coexist and share devices
Demand Based Switching	Dynamically transition one or more CPU cores/threads to low power state and/or lowering frequency based on workload to improve power consumption
Fully Buffered DIMM (FBD)	High speed serial interface with DDR2 DRAM technology to increase DIMM capacity and reduce board routing congestion.

© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Read with remote modified line transaction flow





© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Reliability, Availability and Serviceability (RAS) Features

- Single Device Disable Code (SDDC) / Server ECC (SECC) allows x4/x8 DRAM device failure recovery
 - For x4/x8 devices:
 - Corrects single and detects double device errors
- Scrubbing – correcting error in memory
 - Patrol scrubbing periodically corrects errors
 - Demand scrubbing writes correct data to DRAM on a correctable error
- DIMM sparing - “spare” DIMM to replace a failing rank
- Memory Mirroring – mirrored branches of memory
 - On an uncorrectable error , retry from other redundant branch
- FSB address and data parity error checks
- FBD channel retry on CRC error
- 8B/10B and CRC-32/16 on PCI Express interfaces
- DM parity error protection, ECC on SF
- Data poisoning

© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Bensley Performance



© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Blackford Northbridge (BNB) Uarch Perf Features

- Speculative read to memory launched
- Early initiation of defer reply reduces idle latency
- Defer replies, cross-bus snoops, IO-snoops, & processor requests arbitrate for the FSB using a multi-tier weighted scheme
- 16MB SF for improving system performance
 - Eliminates un-necessary cross-bus / IO snoops and decreases bus utilization
- For SF-misses, optional mode to defer memory requests instead of completing in in-order mode
- Weighted arbitration amongst PCI-E ports to balance bandwidth

© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

BNB Uarch Perf Features – Contd.

- Uses FBD protocol capabilities for higher memory channel efficiency
- Higher reordering capabilities than prior generation MC
- Intelligent rd/wr switching schedules writes while conflicts are present on reads or vice versa
- Supports page-close policy for highly threaded applications
- Supports DDR2 “posted CAS” feature for higher memory channel efficiency

© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Latencies and Bandwidth

Performance Metric	2004 DP (LH)	2006 DP (BF)	LH vs. BF
FSB Peak Bandwidth	6.4 GB/s	21 GB/s	3x+
FSB Sustained Bandwidth*	<3.2 GB/s	9+ GB/s	2.8x+
Memory Peak Bandwidth	6.4 GB/s	21 GB/s	3x+
Memory capacity	16 GB	64 GB	4x
Memory idle Latency **	85 ns	87 – 102 ns	1x – 1.17x
Average Memory loaded latency for TPC-C traffic mix	180 – 200 ns	115 – 125 ns	0.6X - 0.65x
Delivered Bandwidth for Concurrent FSB & IO traffic	<3.2GB/s	12 GB/s	3.75x+

Blackford (BF) delivers significant boost in bandwidths, higher memory capacity and lower loaded latency

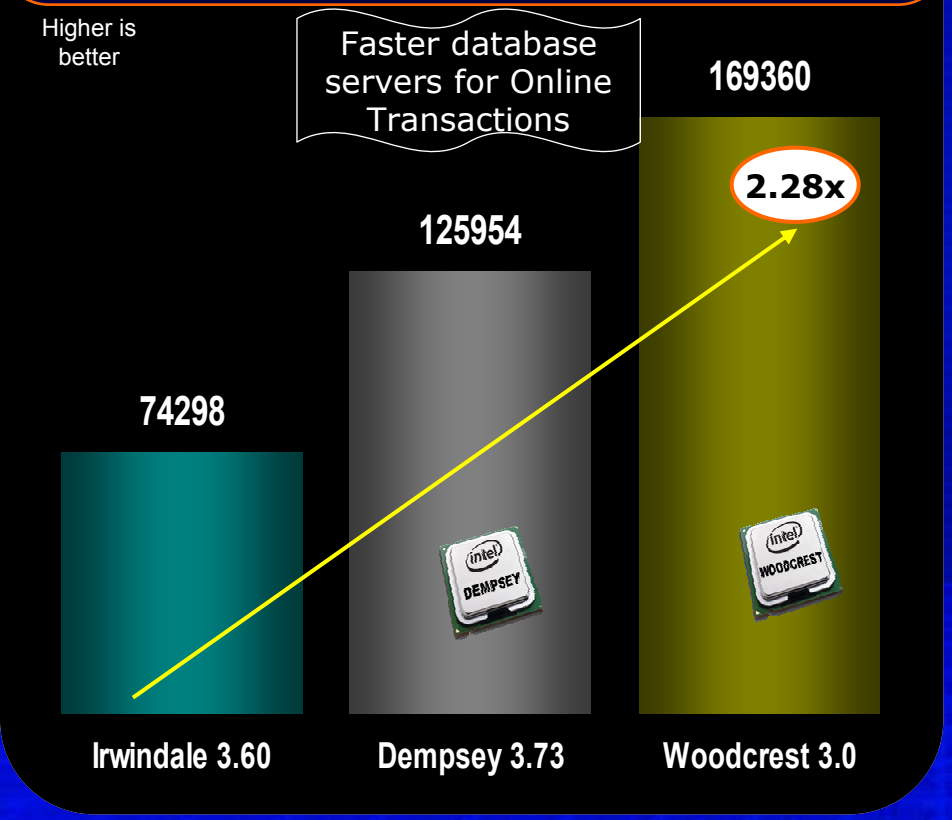
* Only FSB-to-memory traffic with no IO

** 1333 MT/s, 4xFBD-667 (5-5-5)

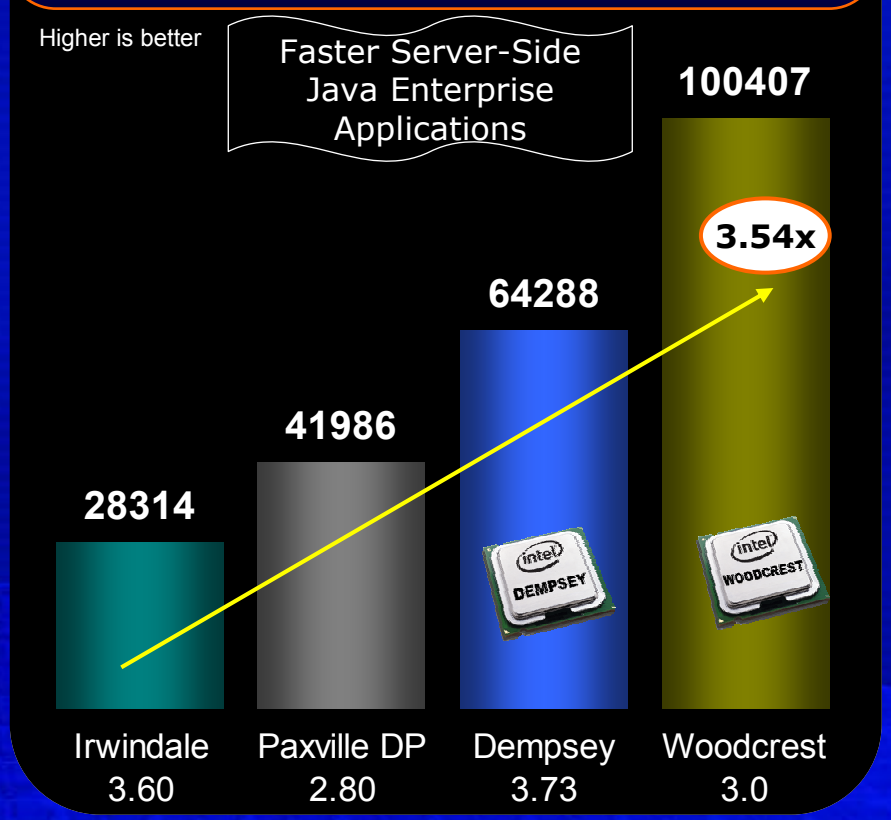
© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Dual-Core Intel® Xeon® processor 5000 & 5100 series based Servers
Data base Server and Java Performance

TPC-C* with Microsoft SQL Server



SPECjbb2005*

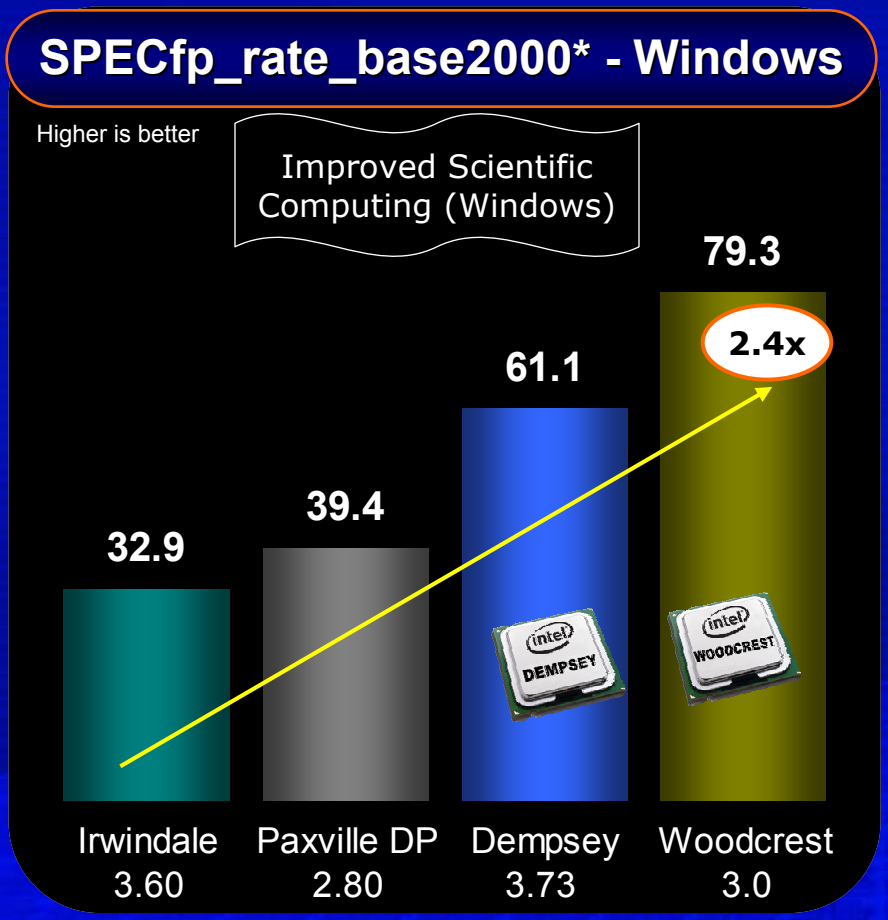


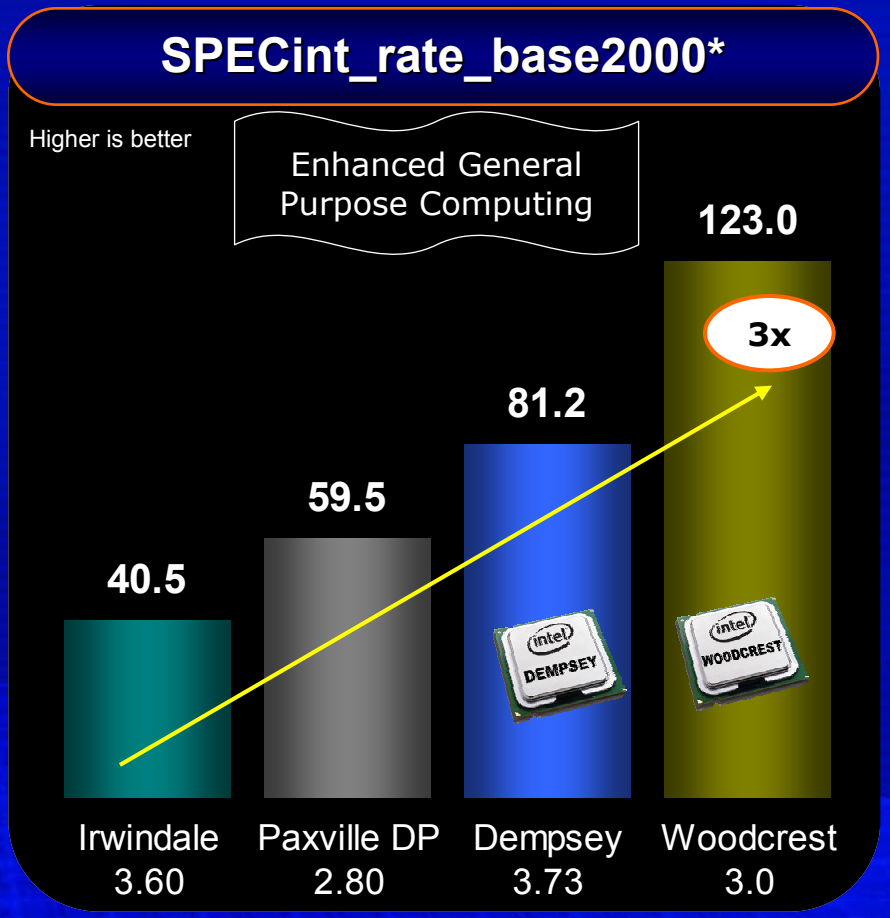
- Key details: Microsoft SQL2005 database, 64-bit Software stack (OS and database), Irwindale w/ 16 GB mem , Dempsey w/ 32 GB mem and Woodcrest w/ 64 GB mem

Record-setting TPC-C delivers significant performance advantage

© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Dual-Core Intel® Xeon® processor 5000 & 5100 series based Servers
Floating Point and Integer Performance

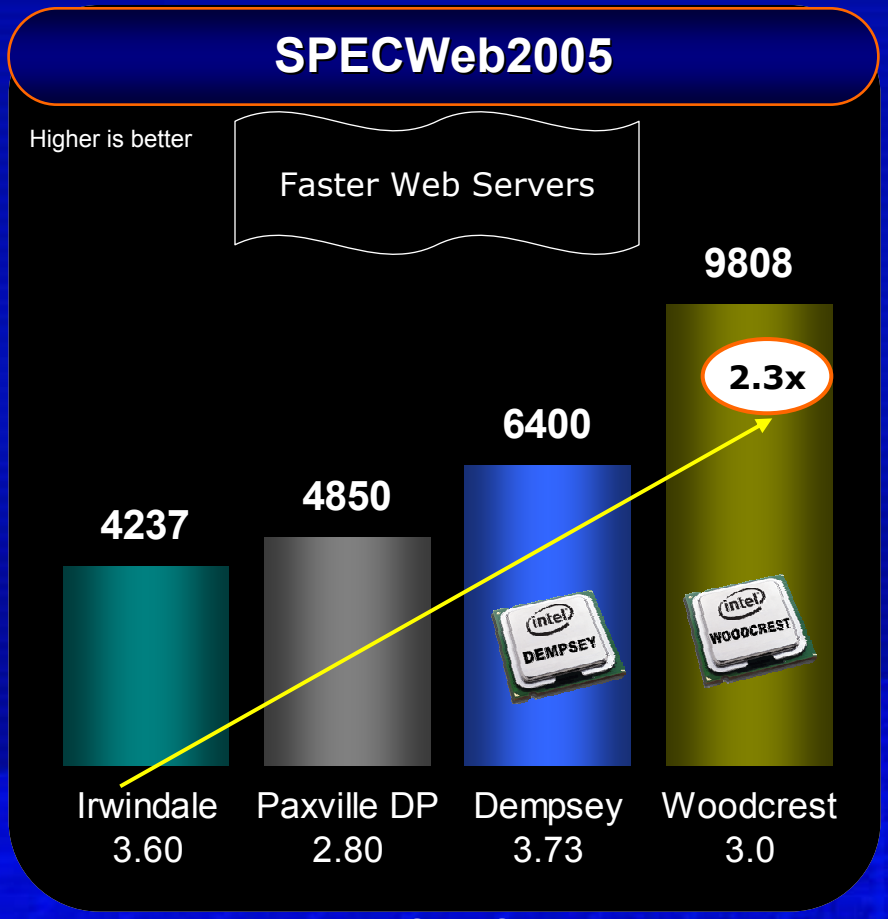


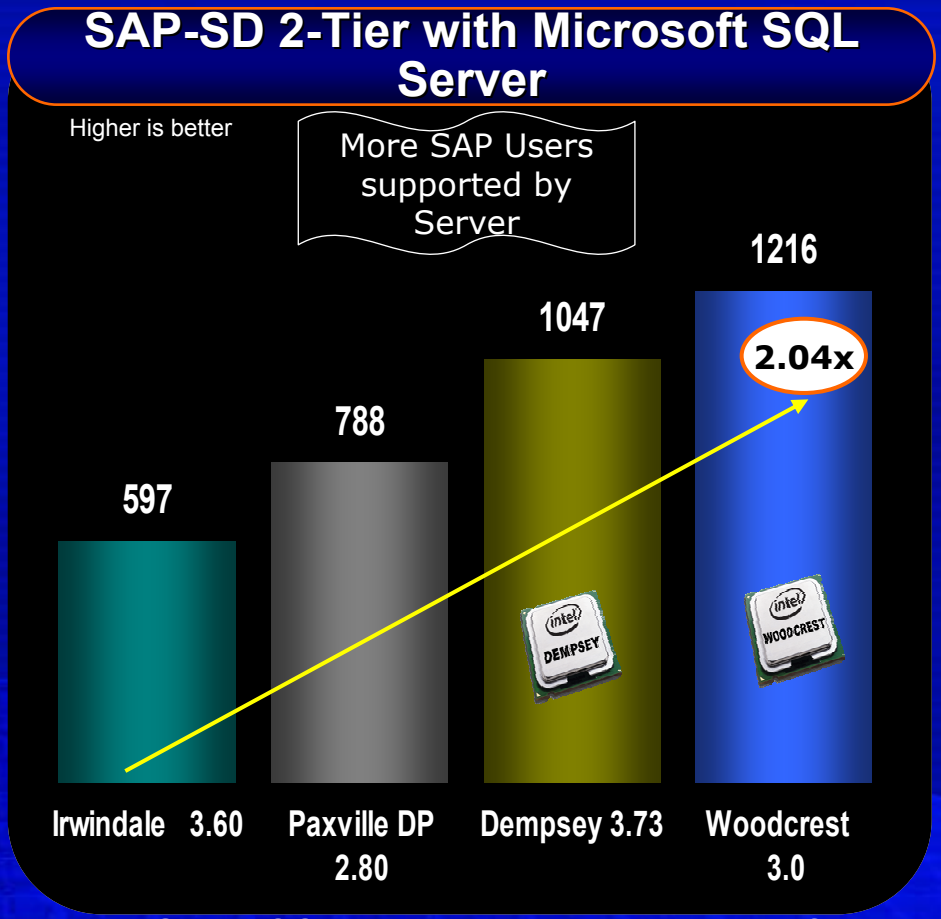


Significant Performance Boost in general purpose and scientific computing applications

© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Dual-Core Intel® Xeon® processor 5000 & 5100 series based Servers
Enterprise Resource Planning and Web Server Performance



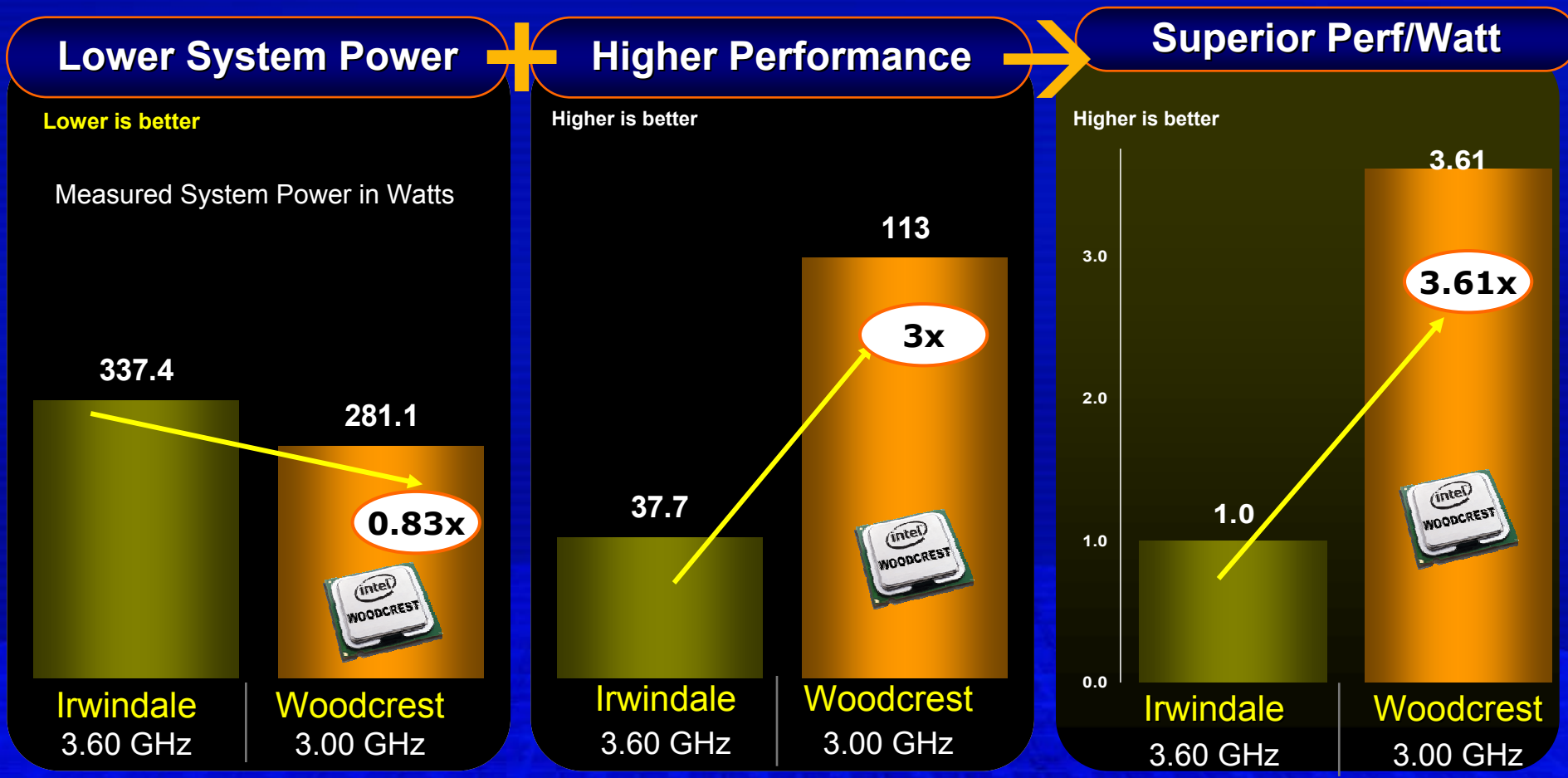


- Key details: Two-tier SAP Sales and Distribution benchmark, SAP ECC Release 5.0 (64-bit), Microsoft Windows Server 2003 Ent. Edition (64-bit), SQL Server 2005 Database (64-bit)
- **First DP Server Platform to cross 1000 SD users**

Significant gains on Web Server and ERP Performance

© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

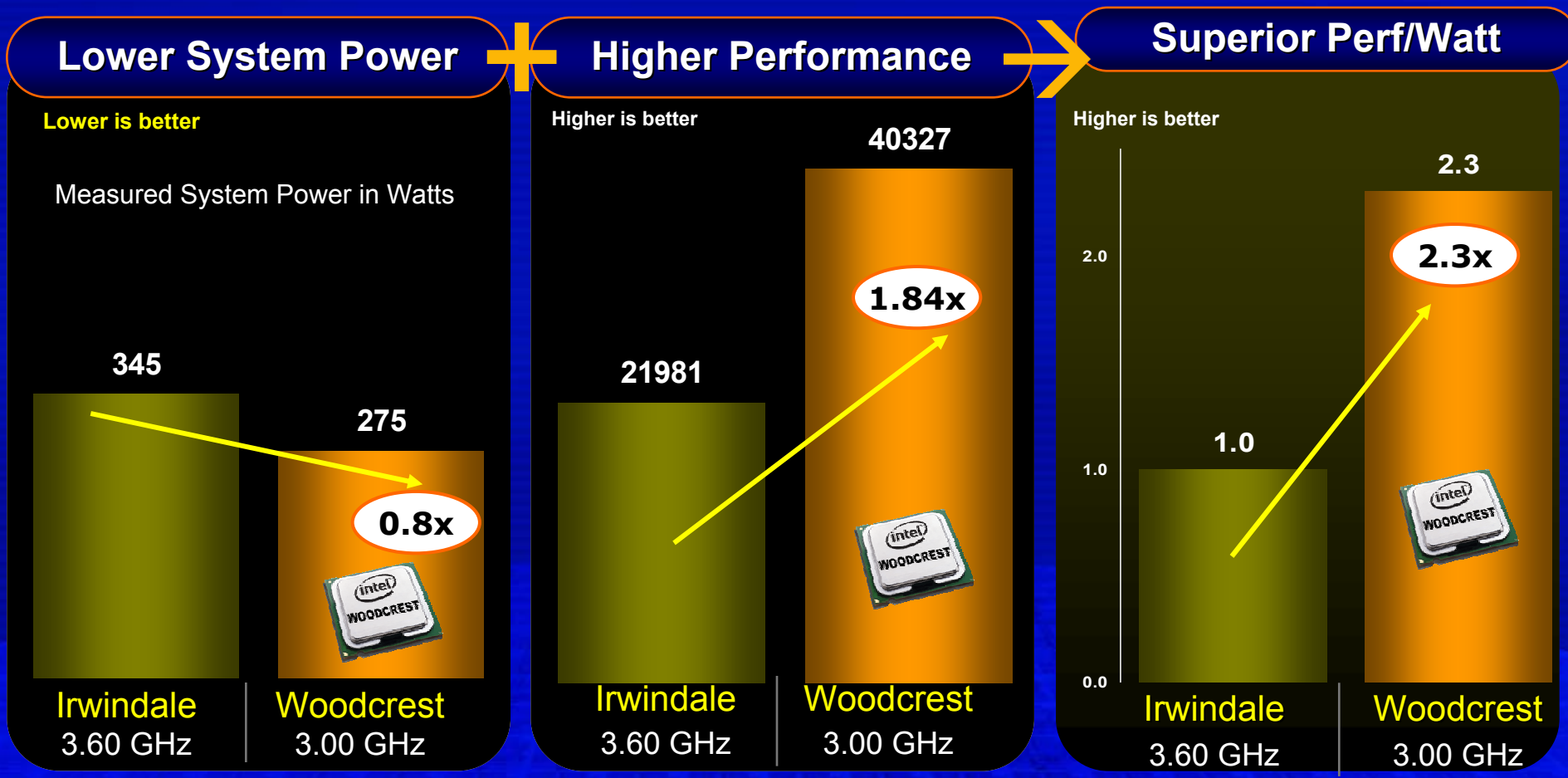
Dual-Core Intel® Xeon® processor 5100 series based Servers
Energy Efficient General Purpose Computing with SPECint_rate_base2000*



3.6x better Performance/Watt on SPECint_rate_base2000

© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Dual-Core Intel® Xeon® processor 5100 series based Servers
Energy Efficient web serving with WebBench* 5.0



2.3x better Performance/Watt on WebBench

© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Summary

- Bensley provides technology *leadership* in DP servers/workstations
 - Architected for Dual and Quad Core Xeon CPUs
 - *Dual Independent Buses* allow for faster system bus speeds
 - FBD technology
 - Increased IO connectivity/throughput
 - Intel® I/OAT and TPM v1.2 Support
 - Enhanced RAS / Debug/DFT features
- Greater platform dependability, performance and increased value to enterprise front end, small to medium businesses
- For more information on the performance of Intel products, visit *http://www.intel.com/performance/resources/limits.htm*

© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Acknowledgment

- Jimbo Alexander, David C Lee, Mike Wiznerowicz, Sin Tan, Sivakumar Kuppuswamy, Perry Taylor, Mark Swanson, Sundar Iyengar, Vish Viswanathan, Bruce Christenson, Jeff Wilder, Chitra Natarajan, Suresh Chittor, Suneeta Sah, Rami Naqib, Rajesh Pamujula, Rajat Agarwal, Chih-Cheh Chen, Chris Van Beek, Debendra Das Sharma, Rajee Ram, Amir Taraghi, Dominic Gasbarro, Brian Parris, Byron Sonner, Subba Vanka

- Bensley Design, Circuit, Platform, Processor, Memory, Performance Modeling, Manufacturing, Validation, Planning, Enterprise Performance Marketing and Management Teams

© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Additional Legal Disclaimer

- Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Buyers should consult other sources of information to evaluate the performance of systems or components they are considering purchasing. For more information on performance tests and on the performance of Intel products, visit http://www.intel.com/performance/resources/limits.htm or call (U.S.) 1-800-628-8686 or 1-916-356-3104.

- All dates and products specified are for planning purposes only and are subject to change without notice

- Relative performance for each benchmark is calculated by taking the actual benchmark result for the first platform tested and assigning it a value of 1.0 as a baseline. Relative performance for the remaining platforms tested was calculated by dividing the actual benchmark result for the baseline platform into each of the specific benchmark results of each of the other platforms and assigning them a relative performance number that correlates with the performance improvements reported.

- 64-bit Intel® XeonTM processors with Intel® 64 requires a computer system with a processor, chipset, BIOS, OS, device drivers and applications enabled for Intel ® 64 . Processor will not operate (including 32-bit operation) without an Intel® 64-enabled BIOS. Performance will vary depending on your hardware and software configurations. Intel® 64-enabled OS, BIOS, device drivers and applications may not be available. Check with your vendor for more information.

- SPECint2000 and SPECfp2000 benchmark tests reflect the performance of the microprocessor, memory architecture and compiler of a computer system on compute-intensive, 32-bit applications. SPEC benchmark tests results for Intel microprocessors are determined using particular, well-configured systems. These results may or may not reflect the relative performance of Intel microprocessor in systems with different hardware or software designs or configurations (including compilers). Buyers should consult other sources of information, including system benchmarks; to evaluate the performance of systems they are considering purchasing.

© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Backup



© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Glossary/Acronyms

- MT/s – Million Transfers per second
- x4, x8, x16
 - Refers to the PCI-Express link width per direction defining the number of lanes in the interface. Bandwidth doubles for each configuration. x4 represents 1 GB/s, x8 has 2 GB/s and x16 delivers 4 GB/s
- 2 and 3 Load FSB
 - Refer to the number of sockets in the FSB including the North bridge.
 - A single socket (dual or multicore) CPU with a North Bridge is considered a “2 Load device” while 2 sockets plus North bridge is a 3 load device. Electrical parameters such as trace length, frequency, voltage ramp, reflections and I/O drivers place a limit on the number of sockets for a given frequency.

© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Glossary (contd)

- AMB – Advanced Memory Buffer.
 - This is the component onboard the FB-DIMM that converts data between the point-to-point FBD interface and the DDR2 interface. On the FBD interface, it unconditionally forwards all packets both northbound and southbound. On southbound packets, it checks to see whether a command or data is directed at it. The AMB does not examine or check integrity of any northbound data passing through it from other downstream AMBs; it merely retransmits the data northbound.
- ECC – Error correcting code.
 - Additional DRAMs are used to contain check-bits that are calculated for each 64 or 128 bit data group. The check bits stored in the DRAMs are used to automatically correct some types of data errors. DIMMs only store the ECC check bits; all error correcting and checking are done in the northbridge. ECC is primarily used to protect the integrity of DRAM memory cells.
- Correctable Error – a data error that is corrected.
 - These errors are corrected by ECC logic in the Bensley Northbridge, based on ECC check-bits stored in the DIMMs. The northbridge supports a feature called "Demand Scrubbing". This feature, when enabled, will cause the corrected READ data to be automatically written back into the DIMMs.
- CRC – "Cyclic Redundancy Check".
 - CRC is a checksum calculated across a block of data transfers. The sender calculates the CRC based on a simple XOR polynomial calculation in hardware; it sends the CRC value at the end of the block transfer. The receiver also calculates a CRC on the incoming data and that value is checked against the incoming CRC value sent. In FB-DIMM protocol, a CRC is calculated on **most* packets sent northbound (towards the northbridge), and all packets southbound (towards the DIMMs). The AMB on the DIMMs is responsible for checking the CRC of southbound packets, and calculating a CRC to send northbound.

© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Glossary (contd)

- Scrubbing –
 - Process of correcting a correctable memory location that may be caused by soft / alpha errors on the FBD. There are two types of scrubbing– demand and patrol.
 - Demand Scrub
 - If the CPU or I/O makes a demand read and the read data from memory turns to be a “correctable” ECC, it is corrected and data sent to source as well as memory is updated.
 - Patrol –
 - This is a background activity initiated by the northbridge to seek out and fix memory errors. Patrol Scrub scans all of memory doing simulated “READs” while checking for ECC errors. If any ECC errors are detected during this process, they are logged as Patrol errors. Correctable errors are corrected and written back into memory.

© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

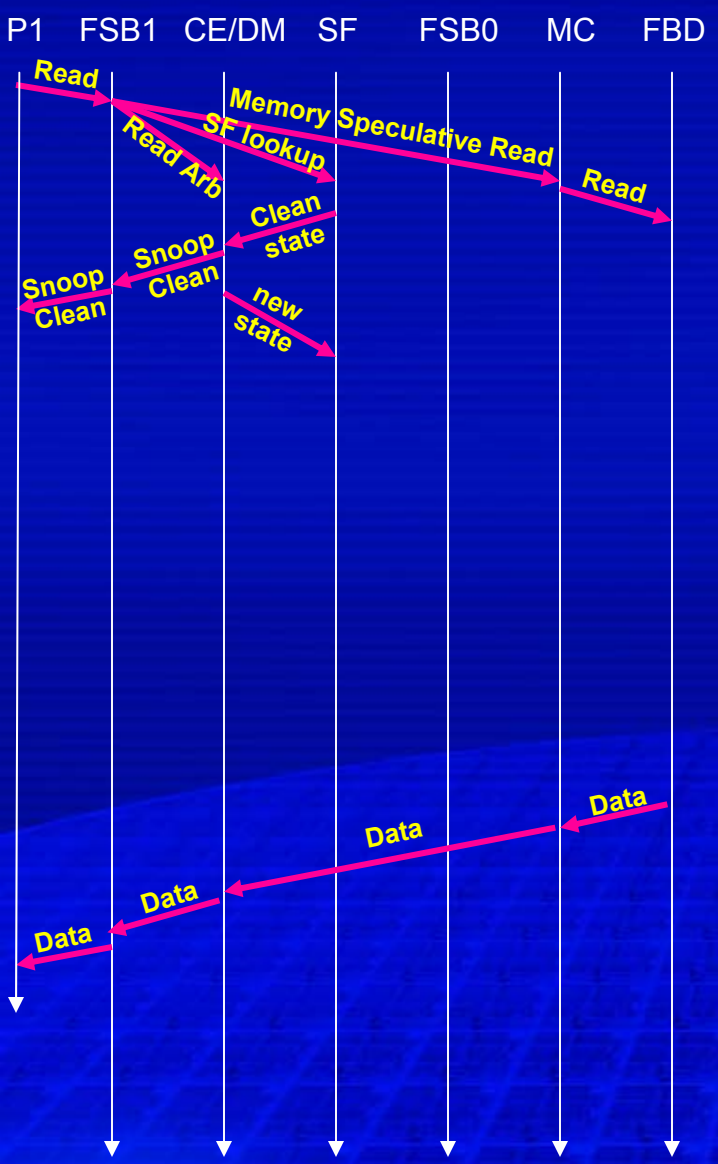
LH (2004-DP) vs. BF (2006-DP) Idle Latencies

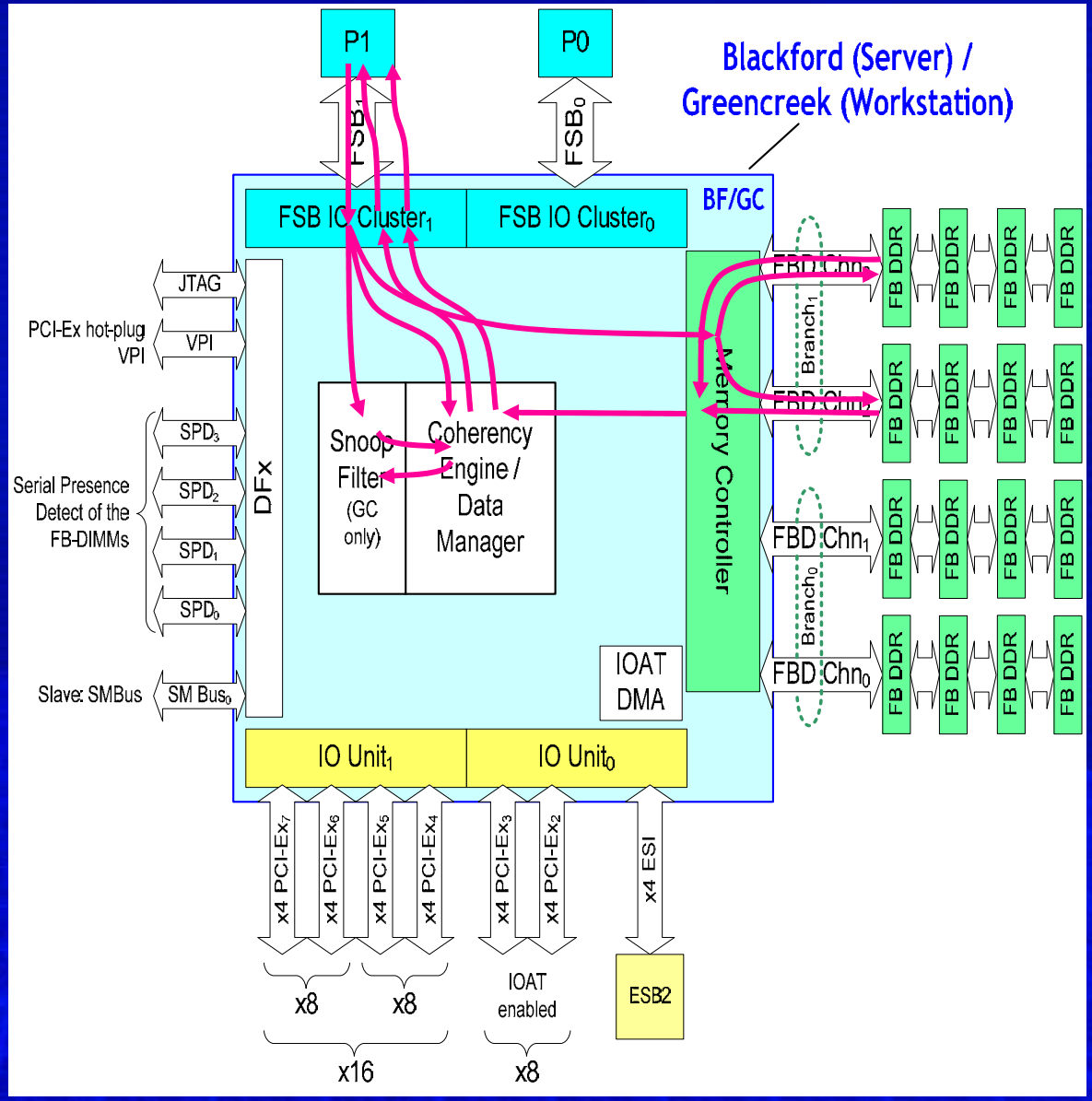
Configuration		Memory read latency	Remote HITM latency	Local HITM latency
1333 MT/s / FB-DDRII-667 (5-5-5)	1 DPC	87	69	30
	2 DPC	93		
	4 DPC	102		
Lindenhurst 800 MT/s / DDRII-400 (3-3-3)		85	N/A	40

ADS – Address/request strobe on the FSB
Idle latencies higher than Lindenhurst primarily due to FBD
However, significantly lower average loaded latencies due to higher FSB and memory bandwidth than Lindenhurst platform

© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

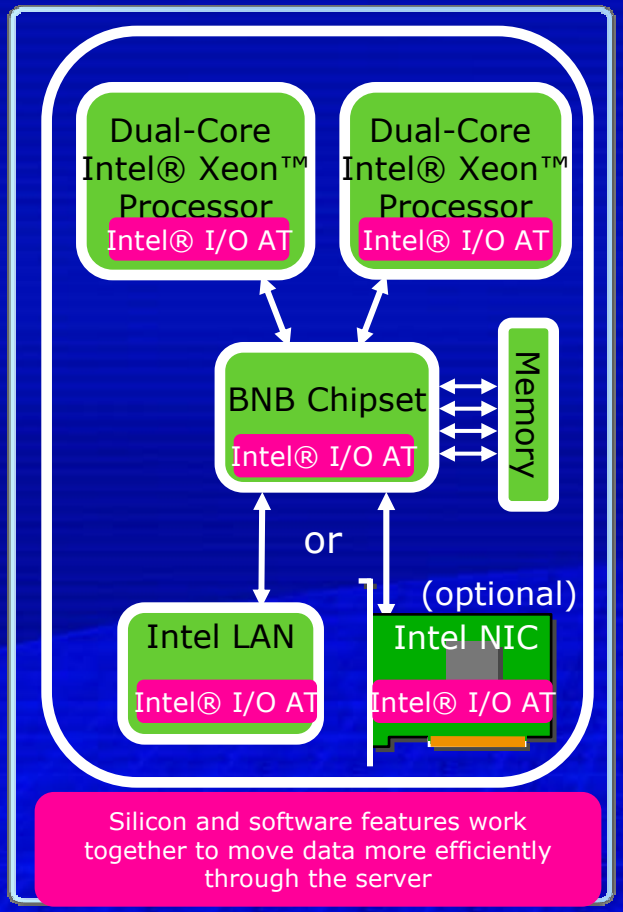
Memory read with clean snoop transaction flow





© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Intel® I/O Acceleration Technology (I/OAT)



Fast

Faster data movement - 2x better max. data throughput*, relieves CPU from performing mundane data transfers (copy)

Scalable

Intel® I/OAT scales seamlessly up to 8 GbE ports
I/O perf. increases with CPU improvements

Reliable

Uses the trusted Windows & Linux TCP/IP stacks
Preserves existing LAN features

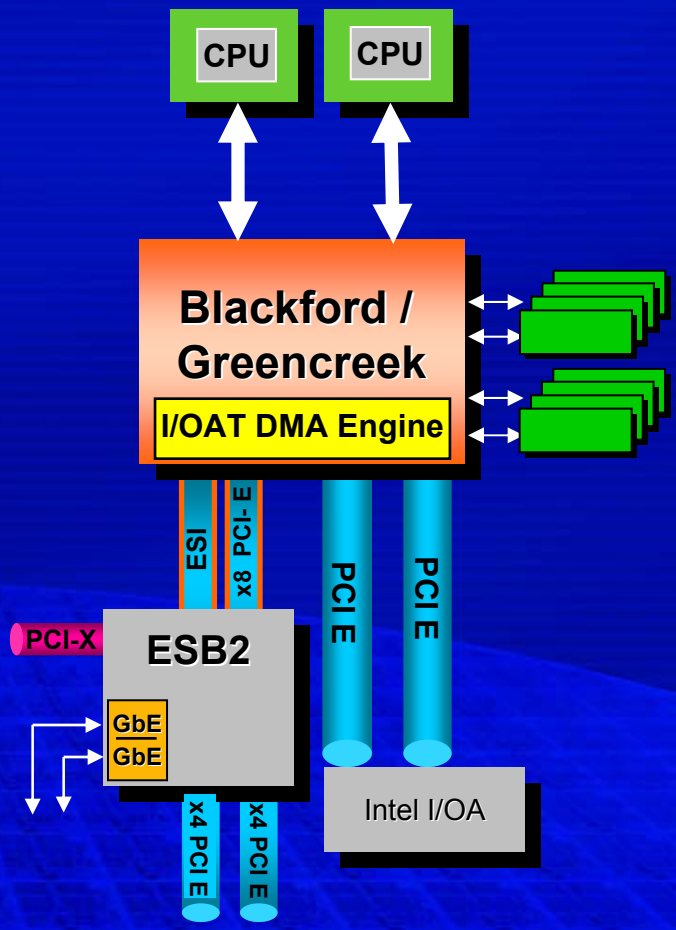
* comparisons are using 6 port Linux Configuration, comparing to previous Q2'05 64-bit Intel Xeon based servers prior generation

Intel® I/O Acceleration Technology requires use of Dual-Core Intel® Xeon Processor 5000 Sequence Processors, Intel® 5000 Sequence Chipsets, Intel® 6321 ESB I/O Controller Hub, either Intel® 82563EB/82564EB or Intel® PCIe Server Adapter with Intel's Nyssa 4.1 Beta Release or later, Microsoft Server 2003 with Scalable Network Pack or Linux 2.6.12 Kernel (or later)

© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.

Intel® I/O Acceleration Technology

- BNB supports 4 Channel DMA Engine
- Memory-Memory, Memory-MMIO transfer through onboard DMA engine
- Descriptor access mechanism for Source/Destination addresses with maximum transfer length of 4KB/1MB
- Automatic Scatter/Gather operation from source to destination memory
 - i.e. data moved from different locations to a single contiguous space or vice versa based on descriptor pointer
- splicing and appending operations to Descriptor link list
- Byte Alignment
- DMA BAR/MMIO registers
- Interrupt signaling (MSI or legacy)
- Status update to device to signal end of operation
- Error handling / reporting



© Copyright 2006, Intel Corporation. All rights reserved.
*Third party marks and brands are the property of their respective owners.